# Automatic Detection And Categorization Of Skin Lesion For Early Diagnosis Of Skin Cancer Using Yolo-V3 - Dcnn Architecture

Akila Victor[1]*, S. Oswalt Manoj[2], R. Rama Abirami[3], Monika Arya[4]

[1]School of Computer Science and Engineering,

Vellore Institute of Technology, Vellore, Tamil Nadu, India.

[2]Department of Computer Science and Business Systems,

Sri Krishna College of Engineering and Technology,

Coimbatore, Tamilnadu, India.

[3]Department of Computer Science and Engineering,

Sri Krishna College of Engineering and Technology,

Coimbatore, Tamilnadu, India.

[4]Department of Computer Science and Engineering

Bhilai Institute of Technology, Durg C.G.

*akilavictor123@gmail.com

**Abstract:** Skin cancer was among the most lethal illnesses in the globe, especially malignant melanoma, benign seems to be a kind of skin cancer that would be the most deadly because of its rapid development, affecting a huge number of individuals globally. Early diagnosis of skin cancer seems to be critical, hence our research employs the YOLO v3 - DCNN architecture to discover and categorize the deadliest kinds of skin cancer. Initially, YOLO v3 generates the feature map, simultaneously color features are extracted using color moments with QuadHistogram, whereas Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform(RCT) is generated texture features, and both (color and texture) features are get fused. Then, fused features are fed into the Deep Convolutional Neural Network (DCNN) which classifies the different types of skin cancer. Finally, our proposed approach is compared with the existing works. As a consequence, when contrasted to the baseline techniques, our proposed YOLO-v3 –DCNN has a greater accuracy.

**Keywords:** Skin cancer, Deep Convolutional Neural Network (DCNN), Deep Learning, You Only Look Once (YOLO-v3).

## INTRODUCTION

The skin seems to be the body's largest organ, and it shields it from temperature, radiation, as well as illness. This also aids in body temperature control including fat as well as water storage. Among the most drastic concerns about the skin in the body has been the threat of skin cancer infection [1]. Skin cancer is really an infection generated by irregular cells

which affects humans. Skin cancer starts with cells, which have been the fundamental components of the skin; skin cells proliferate as well as split to produce new cells. Skin cells age and die on a daily basis, and new ones replace them. This meticulous approach might lead to errors from time to time. Whenever the skin does not need the them, fresh cells develop, while old cells perish when they don't. Excessive cells clump together to form a mass of tissue termed as a tumor [2, 3]. If this is not diagnosed early, it may have the ability to spread to certain other areas of the body [4].

Skin cancer has been typically categorized as melanoma and nonmelanoma [5]. Melanoma seems to be a severe, uncommon, as well as lethal kind of skin cancer. As per the American Cancer Society, melanoma skin cancer accounts for just 1% among all occurrences, but it does have a higher mortality rate [6]. Melanoma grows from melanocytes, which are skin cells. It begins whenever healthy melanocytes multiply excessively, results in formation of a malignant tumour. This will have the potential to adversely affect any part of the human body. It is especially common on sun-exposed areas such as the hands, face, neck, as well as lips. Melanocytes are already responsible for the production of dark colours on the skin, hair, eyes, and other body parts. As a consequence, most melanoma tumours appeared brown or black in colour. Nonetheless, in a small percentage of instances, melanomas do not establish colour and instead look pink, red, or purple [7].

Melanoma cancers should only be cured if they have been identified early; or else, they spread to the rest of the body eventually kill the sufferer [8]. Nonmelanoma tumours have become less complicated to cure than melanoma tumours. Melanoma research [9] was among the first to emphasis on Computer-Aided Diagnosis (CAD) [10, 11] via applying automated strategies for the earlier identification of all skin lesions. Such systems utilize classical computer vision technology to extract attributes like shape, colour, as well as texture to supply a classification model [12-16].

Early identification of melanoma has been seen in studies to considerably lower the mortality percentage from melanoma cancer [17]. The fact that earlier identification of melanoma remains tough, even for specialists, is indeed a major issue. As a result, utilizing a strategy to simplify the diagnosis might be beneficial to professionals [18–21]. Because of the wide variety of skin lesions seen in dermatology, an automated skin cancer diagnosis is a difficult undertaking. Various automated approaches for detecting and diagnosing skin cancer have been developed throughout the years.

The traditional way of medical picture analysis involves a succession of low-level pixel processing technologies. Because of the poor distinction between both the surrounding skin as well as the lesion region, accurate segmentation of the lesion is also difficult. . Recent advancements in this area shown that deep learning has been the most beneficial machine

learning approach applied towards the issue and utilized in a variety of applications including speech recognition [22], pattern recognition [23], as well as bioinformatics [24]. Deep learning methodologies had produced amazing outcomes in several areas when compared to other traditional approaches to machine learning. Numerous deep learning methodologies also being investigated in latest years for computer-dependent skin cancer screening. Using these strategies speeds up the diagnostic procedure but also lessens human errors. This can affect the effectiveness and quickness with which physicians as well as radiologists identify melanoma. Even though the CNN design produces superior results, there are significant accuracy limitations in extracting features from skin lesions, necessitating the development of a new deep learning-dependent architecture. As a consequence, identifying skin cancer at an earlier phase, our research proposed YOLO v3 – DCNN architecture. This research can be divided into the following stages:

- Initially, You Only Look Once (YOLO)-v3 architecture uses a combination of YOLO v2, Darknet-53, and Residual networks to extract features from skin lesions. Simultaneously, color-texture characteristics were retrieved by Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform(RCT) and color moments with QuadHistogram. Then extracted features are get fused.
- In addition, the fused characteristics were fed into the DCNN for categorization of skin lesions including malignant, benign, as well as basal cell carcinoma (bcc).

The remainder portions of the paper were assembled as follows: Section 2 gives a review regarding feature extraction as well as categorization of skin cancer. Section 3 explains the proposed YOLO v3 - DCNN methodology. Section 4, elaborates on the experimental results, comparisons, and discussions, followed by Section 5, concludes the paper.

## LITERATURE SURVEY

Much study has been undertaken in latest days employing various artificial neural networks for early diagnosis of skin malignancies, which is addressed below:

Xie et al [25] using digital dermoscopy images, established a unique approach for categorizing melanocytic tumors as benign or malignant. The suggested border characteristics discussed here were supposed to be insensitive to lesion object incompleteness. The suggested system functioned in three steps. Initially, a self-generating NN was being used to retrieve lesions from photographs.The second phase retrieved characteristics such as tumor boundary, texture, and color information. Feature dimensionality reduction was utilized to remove less relevant or noisy features, resulting in improved classification performance. Subsequently, lesions were categorized utilizing a NN ensemble system in

the final phase. Ensemble NN optimizes categorization effectiveness by incorporating back propagation (BP) with fuzzy neural networks.

Yu et al. [26] offer a novel melanoma detection strategy depend on extremely deep convolutional neural networks (CNNs). We seamlessly merge the two phases and provide an automated framework that does not require manual intervention. A fully convolutional residual network (FCRN) containing 16 residual blocks is being used in the segmentation step to optimize effectiveness. During categorization, the chosen strategy would use an average of SVM plus softmax models. That has an 85.5 % accuracy for melanoma categorization using segmentation as well as an 82.8 % accuracy without segmentation. Among the most serious barriers in deploying deep CNNs with medical image processing tasks is indeed a scarcity of high-quality training set.

Dorj et al. [27] established a strategy for categorizing skin lesion photographs into four distinct groups. To satisfy the necessity for an intelligent as well as rapid skin cancer categorization, a pre-trained deep CNN termed AlexNet is being used for extraction of features, accompanied by an error-correcting output coding SVM functioning as a classifier.

Nida et al. [28] suggested an innovative methodology centered on RCNN as well as FCM clustering enabling accurate & automated Melanoma area segmentation among dermoscopic pictures. A deep Region-dependent convolutional neural network (RCNN) effectively recognizes the numerous afflicted areas mostly in form of bounding boxes utilizing Fuzzy C-mean (FCM) grouping, that facilitates localization.

Adegun et al [29] introduced a unique paradigm for automated skin cancer detection which includes segmentation as well as categorization of skin lesions. A deep learning-dependent CAD architecture comprised of a multi-scale encoder-decoder segmentation system as well as an FCN-dependent DenseNet categorization network has also been introduced for the recognition as well as categorization of skin lesion photos in order to recognise skin cancer illnesses.

Akram et al. [30] suggested an unique methodology for skin lesion categorization which utilizes deep feature data to create the far more discriminant feature vector whilst preserving the initial feature space. Following fine-tuning, they incorporated the obtained data from the chosen pre-trained systems, that considerably elevated categorization accuracy. In the recommend approach, the researchers utilized less than 3% of the total characteristics, that not only enhances categorization accuracy by eliminating redundancies but also minimizes computing time. After execution of this approach, researchers make the following claims: (a) fusion of retrieved characteristics from a collection of pre-trained systems promotes overall accuracy; while (b) the inclusion of an attribute selection plus dimensionality reduction phase enhances categorization outcomes greatly.

Khan et al [31] used a unique technique that not only identifies skin cancer but also classifies it. The suggested technique is basically focused on saliency assessment as well as the identification among the most discriminating deep characteristics. To enhance lesion distinction, the suggested Gaussian system is employed, followed by a colour space transfer from RGB to HSV. By keeping that foreground and background as distinct as possible, the new color space aids in the building of saliency maps that use inner and outer discontinuous windows. The inception CNN architecture on two basic output tiers is being used to capture deep information from segmented photos. Those retrieved characteristics are then fused utilizing suggested decision-controlled parallel fusion approach before even being chosen employing the suggested window distance-controlled entropy characteristics selection technique. The most differentiating characteristics were then put to categorization.

Khan et al. [32] developed a novel strategy for skin cancer localization as well as diagnosis by merging a deep learning system with an iteration-controlled Newton-Raphson (IcNR) oriented attribute selection approach. This suggested system comprises of three fundamental steps: lesion localization utilizing a faster region-dependent convolutional neural network (RCNN), deep characteristic retrieval, as well as feature selection that use the IcNR approach. There in localization step, a novel contrasted stretching strategy centered on the bee colony method (ABC) has been applied. The augmented images, together along with underlying ground truths, were then fed into Fast-RCNN, resulting yields segmented pictures. A pre-trained model, DenseNet201, has been utilised to retrieve deep features via transfer learning, which would then be exposed to a decision stage through the suggested IcNR approach. Ultimately, the most discriminating characteristics are categorized employing layered feedforward neural networks.

Khan et al. [33] developed a totally automated computer aided diagnostic (CAD) solution relying on the deep learning framework. The initial dermoscopic images were once pre-processed employing the decorrelation formulation strategy, and also resulting images are then supplied to the MASK-RCNN for lesion segmentation. This MASK RCNN network is again trained utilizing the segmented RGB pictures obtained from the ground truth images. Incorporating a GrabCut algorithm as well as an Adaptive Neuro-Fuzzy Classifier (ANFC) framework.

Sikkandar et al [34] propose a novel segmentation-dependent categorization framework for skin lesion identification. Initially, a Top hat filter as well as inpainting techniques were used in the preprocessing stage. The preprocessed pictures are then segmented using the Grabcut method. The feature extraction approach is then utilized in collaboration with a deep learning-dependent Inception model. Ultimately, the dermoscopic pictures were categorized using an adaptive neuro-fuzzy classifier (ANFC) network.

Saeed et al. [35] proposed an advanced skin cancer detection and classification method based on DCNN architectural variations. Transfer learning, fine-tuning, the ensemble technique, data creation, and augmentation are all viable methods for reducing labeled data insufficiency, which is prone to overfitting. It also helps to improve the overall effectiveness of skin lesion categorization in CAD systems.

However, no research has been conducted to increase the accuracy and speed of the process by combining artificial intelligence approaches and traditional feature extraction methods with YOLO-based methodologies.

## AUTOMATIC DETECTION FOR EARLY DIAGNOSIS OF SKIN CANCER AND CLASSIFICATION USING YOLO V3 – DCNN ARCHITECTURE

Skin cancer is among the most lethal diseases in the globe, therefore early identification as well as categorization are necessary. Thus, this paper proposed a combined architecture YOLO version 3 and deep CNN has been presented for detecting and classifying the skin lesion. It explored a You Only Look Once (YOLO)-v3 architecture, which uses a combination of YOLO v2, Darknet-53, and Residual networks to extract features from skin lesions. The darknet's 53 layers are layered with 53 additional layers for the detection head, giving YOLO v3 a total of 106 layers fully convolutional underlying architecture. This layer generates the feature map with high accuracy. Simultaneously texture and color features are extracted in which texture feature extraction can be done employing a Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform(RCT) to provide accurate information about an image's texture, and Color moments are being used to extract colour characteristics with QuadHistogram, wherein the quadtree decomposition being done to the pictures and homogeneous blocks of distinct size are given. Then, the concatenated features have been reduced using the PCA method. The YOLO v3 as well as PCA features are then combined to make accurate features, which would then be fed into a Deep Convolution Neural Network (DCNN) that's been trained with the unique ground truths to increase categorization accuracy. As a consequence, our proposed technique enhances the identification as well as categorization of skin lesions. Fig.,1 depicts the framework of the proposed strategy.

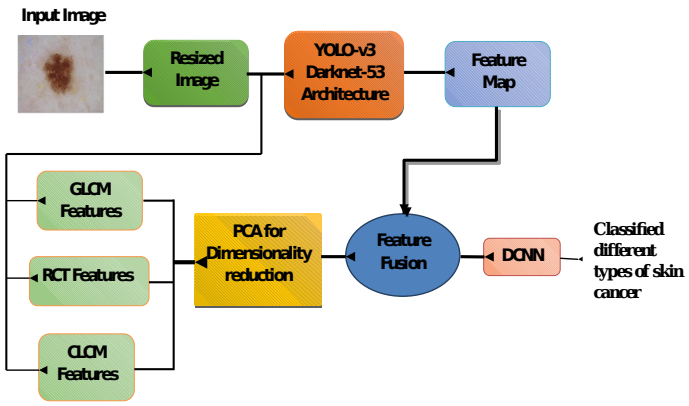


**Fig. 1.** Structure of proposed approach

**Feature extraction by YOLO-v3**

According to recent research,, risk factors could be considerably minimized by detecting it early on, making it almost curable.This prompt identification and categorization necessitate the use of an automated system, even though the method is rather difficult. Thus our research introduces YOLOv3- DCNN architecture to increase the feature extraction as well as categorization accuracy. Firstly, the input data was loaded into the YOLO v3 - DCNN architecture in which features are extracted through the YOLO v3 simultaneously the color and texture features are extracted with Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform and color moments with quad histogram. The obtained features then get fused.

Initially, YOLOv3 uses YOLOv2, Darknet-53, and residual layers, as a novel deep neural network architecture. Darknet-53 has 53 convolutional layers, the majority of which are 3 x 3 and 1 x 1 filters, as well as the darknet's 53 layers, are layered with 53 additional layers for the detection head. Instead of forecasting the locations for bounding boxes from completely linked layers, Darknet leverages anchor boxes. There are 448*448 high-resolution classifiers in our input pictures. The size of the anchor boxes in YOLO v2 is governed by the input picture. It is decreased by a factor of 32.

The network creates five bounding boxes containing five coordinates each: $t_{zx}$, $t_{zy}$, $t_{zw}$, $t_{zh}$, and $t_{zo}$. The width, height, x-coordinate, y-coordinate, as well as class label of every annotation in the collection were expressed by these coordinates.The predictions $b_{zx}$,$b_{zy}$,$b_{zw}$, and $b_{zh}$ are the predicted bounding box coordinates. The cell offsets ($c_{zx}$, $c_{zy}$,) indicate the length of each cell in the picture, as well as the size of the bounding boxes ($p_{zw}$, $p_{zh}$). The predictions are determined as follows:

$$b_{zx}=\sigma¿) +c_{zx} \quad (1)$$

$$b_{zy}=\sigma¿) +c_{zy} \quad (2)$$

$$b_{zw}=p_{zw}e^{t_{zw}} \tag{3}$$

$$b_{zh}=p_{zh}e^{t_{zh}} \tag{4}$$

$$p_{zr}(obj)\times IOU(b_z,obj)=\sigma¿) \tag{5}$$

Where$p_{zr}$ seems to be the conditional class probability and IOU is the intersection of the predicted as well as ground truth bounding boxes, and$\sigma¿$ ) is the change in class label. The input size is reduced by using anchor boxes. The network's last layer is composed of 1000 filters, accompanied by a final layer with a configurable count of filters. The quantity of filters for this layer has been evaluated by the count of classes discovered, as seen below:

$$N_{zfilters}=5\times(5+N_{zclasses}) \tag{6}$$

For $N_{zclasses}$ = 2, the count of filters in the final layer is 35. Whilst YOLO approach extracts characteristics and detects objects using area information, the combination of feature maps containing popular texture as well as colour characteristics was utilized to enhance the algorithm's performance. The techniques for extracting texture and color features are discussed as follows.

The GLCM with Redundant Contourlet Transform (RCT) is used to extract the image's texture features. The GLCM and RCT are widely used and provide accurate information about an image's texture. The texture of a picture is described by GLCM via estimating the frequency of pairs of pixels in the picture with defined values as well as a defined spatial relationship. $p_z(i_z,j_z)$ was utilized in this approach to obtain the image's pixel values at the spatial coordinates $i_z$ ¿ $j_z$. Every measurement generates a 416 * 416 matrix that is utilized to describe each pixel in the picture.

In our research, the Redundant Contourlet Transform (RCT) generates equal-size directional subband pictures. This redundancy gets produced by avoiding any downsampling procedure in the Laplacian pyramid approach; L lowpass approximations of the picture were constructed by employing L suitable low pass filters. When paired with the GLCM, RCT provides high-quality data on the whole texture of the picture. The GLCM and RCT features have a minimum level of complexity and extract higher accurate features.

The colour characteristics are then reconstructed utilising colour consequences utilizing Quad Histogram, in which the images are quadtree

decomposed and homogeneous blocks of varied sizes are provided. Aside from analysing the global texture properties, an orthogonal combination of local binary patterns has also been proposed as a colour predictor for dermoscopic skin lesions.

The color features are extracted color moments with the quad histogram are explained as follows: The implementation of a colour moments-based colour feature extraction approach to the three colour channels, R, G, as well as B, is expanded. In each of the three channels, single as well as multi-integrative co-occurrence matrices were also constructed. As a consequence, a feature vector of dimensions of 208*208*8 be produced for each image. Every feature vector for one single channel approximately 208*208 in size. There seem to be three channels for a single co-occurrence matrix: R, G, as well as B. There have been three extra channels for multi co-occurrence matrices: RG, GB, as well as BR. The original image's single and multi-co-occurrences are then computed as two distinct channels, increasing the dimensions to 8. For proper concatenation, the feature vector is resized to 416*416*2.

Several depths $i_z$of the quad-tree are thus generated, each containing range-blocks of the same size. This research presents some feature histograms. Let $L_z$ denote the greatest depth enforced in the (fractal) decomposition with quad-tree refinement, and $l_z$ denotes the minimum depth. $\Omega_{l_z L_z;k_z}$ *the*domain is defined as:

$$\Omega_{l_z L_z;K_z}=\left\{\left(i_z,j_z\right)\in l_z N_z^2 \middle| l_z\le i_z\le L_z, 1\le j_z\le k_z\right\} \quad (7)$$

Here, $i_z$ is the depth in the quad-tree structure and $k_z$ is the number of feature-bins selected. A function is defined as a histogram $h_z$ on$\Omega_{l_z L_z;k_z}$,

$$h_z:\Omega_{l_z L_z;K_z}\to R_z \; with \; h_z\ge 0 \quad (8)$$

If $\left(i_z,j_z\right)\in\Omega_{l_z L_z;K_z}$ then $h_{zi_z,j_z}=h_z\left(i_z,j_z\right)$is referred to as the value of $h_z$ at$\left(i_z,j_z\right)$. If a histogram $h_z$ on $\Omega_{l_z L_z;K_z}$ satisfies the required extra criteria, it is referred to as a (weighted) quad-tree feature histogram.

$$h_{zi_z,j_z}=w_{zi_z}v_{zi_z,j_z} \quad (9)$$

Where, $\sum_{i_z=l_z}^{L_z} w_{zi_z}\le 1$; $w_{zi_z}\ge 0$and $v_{zi_z,j_z}\ge 0$; $\sum_{j_z=1}^{k_z} v_{zi_z,j_z}=1$, $\forall i_z\in N_z$ with $l_z\le i_z\le L_z$

Equation 9 can be interpreted as follows: It $w_{zi_z}$ can be interpreted to mean weighing the contents of the bins based on depth $i_z$. Also, each depth $i_z$ we have $k_z$ bins$v_{zi_z,j_z}$, the contents of which add up to 1. The color moments with quad histogram color features have a minimum level of complexity and minimum computation power. Finally, color and texture features are obtained through GLCM with RCT and color moments with quad histogram methods.

Then, the texture and color features are merged into a single 416*416*7 matrix and given as input to the PCA which reduces the dimensionality of the concatenated features. The feature map has been generated using the PCA output of 7*7*1024. Moreover, the feature map created by YOLO v3 has 7*7*1024 in size. Both feature maps, including YOLO v3, colour, as well as texture features, are then compressed prior to getting concatenated to establish a fused feature map. This feature map was fed into the DCNN via the feature fusion block, which will then be trained to generate the output.

**Classification based on Deep Convolutional Neural Network (DCNN)**

DCNN is made up of three basic layers: the convolutional layer, the pooling layer, as well as the fully connected layer. Each convolution layer in DCNN has a large count of weights which are subsampled by the pooling layer to give output from the convolution layer whilst lowering the information ratio of the layer below. Ultimately, the outputs of the pooling layer were utilized to inject into the fully linked layers. Convolutional neuron layers, which also include data for a variety of applications such as image categorization as well as numerous 2D matrices, are indeed a key component of DCNN.

This approach may retrieve the regional attributes of the original picture depending on the retrieval of local features. This learning procedure's principal purpose is to build certain kernel matrices for creating more prominent features for the issue of skin cancer identification. In this case, the backpropagation (BP) model was adopted to get the lowest feasible error value for the system. Sliding window-based convolution has been employed for the network. In this research, the rectified linear unit (ReLU) would be employed as the activation function for the neurons via the function $f_z(x_z)=max(x_z,0)$.Max pooling can also be used to reduce the output network size, so that only the optimum levels were assessed as the sliding grid's next layer.

The BP strategy is indeed a gradient descent-dependent system for eliminating neural network error that employs cross-entropy loss as the fitness function. This strategy may be described simply as follows.

$$L_z=\sum_{j_z}^{N_z}\sum_{i_z}^{M_z}-d_{zj_z}^{i_z}\log y_{zj_z}^{i_z} \tag{10}$$

Where $N_z$ would be the sample number, $d_{jz} = \left(0,\ldots,0,\underbrace{1,\ldots,1}_{k_z},0,\ldots,0\right)$ has been the needed output vector, while $y_{zjz}$, seems to be the obtained output vector of the $m_z th$ class, that may be derived utilizing the preceding formula:

$$y_{zj_z}^{i_z}=\frac{e^{f_{zjz}}}{\sum_{i_z=1}^{M_z}e^{f_{zjz}}} \tag{11}$$

The weight penalty is being used in the creation of the function $L_z$ to incorporate a value to enhance the weight values:

$$L_z = \sum_{j_z}^{N_z} \sum_{i_z}^{M_z} - d_{zj_z}^{i_z} \log y_{zj_z}^{i_z} + \frac{1}{2} \eta \sum_{K_z} \sum_{L_z} \omega_{k_z}^2 . l_z \tag{12}$$

Where, $\omega_{k_z}$ indicates the weight of the connection, $L_z$ signifies the total count of layers, while $K_z$ represents the layer $l_z$ connections. This DCNN layer classifies the different types of skin cancers. Figure 2 illustrates the block diagram of DCNN for skin cancer detection.

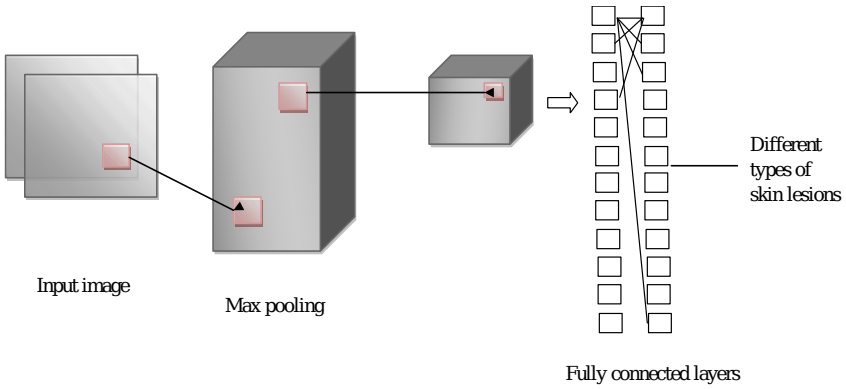


**Fig. 2.** Structure of DCNN in skin cancer detection

As a consequence, the proposed method accurately detects and classifies the different types of skin lesions. Moreover, the next section deals with the implementation results of this proposed strategy.

## RESULT AND DISCUSSION

This section details the implementation findings as well as the performances of our proposed methodology. Also, comparison results of the existing works are presented.

Tool : PYTHON 3

OS : Windows 7 (64 bit)

Processor : Intel Premium

RAM : 8GB RAM

### Dataset Description

The present collection of dermatoscopic pictures is tiny and lacks variety, making it challenging to train neural networks for automatic detection of pigmented skin lesions. The HAM10000 [36] dataset ("Human Against Machine having 10000 training pictures") addresses the aforementioned problem. The authors gathered dermatoscopic pictures from a multitude of populations, which have been recorded and preserved utilizing a diversity of

modalities. Furthermore, the resultant dataset contains 10015 dermatoscopic pictures, which may have been utilized as a training set for educational machine learning. The pigmented lesions comprise actinic keratoses as well as intraepithelial carcinoma / Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, bkl), dermatofibroma (df), melanoma (mel), melanoc (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc). Additionally, the lesion id column in the HAM10000 metadata file might well be utilized to track lesions with a large count of pictures in the dataset.

**Skin lesion detection from the input dataset**

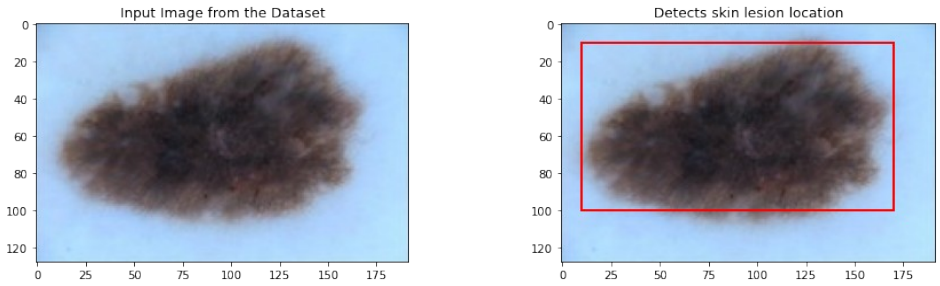


**Fig. 3.** Skin lesion detection from the input dataset

Figure 3 illustrates skin lesion detection, by using our novel technique. The input image from the MNIST HAM10000 dataset is shown in figure 3(a). From the input image (figure 3(a)) the skin lesion is detected using our proposed YOLO-v3 approach, as shown in figure 3(b), in which the red color box is portrayed as a bounding box. The bounding box indicates the region of the lesions for which a high degree of accuracy can be predicted. For the categorization process, all retrieved plus conventional characteristics from the bounding boxes were fused.

**Skin lesion classification**

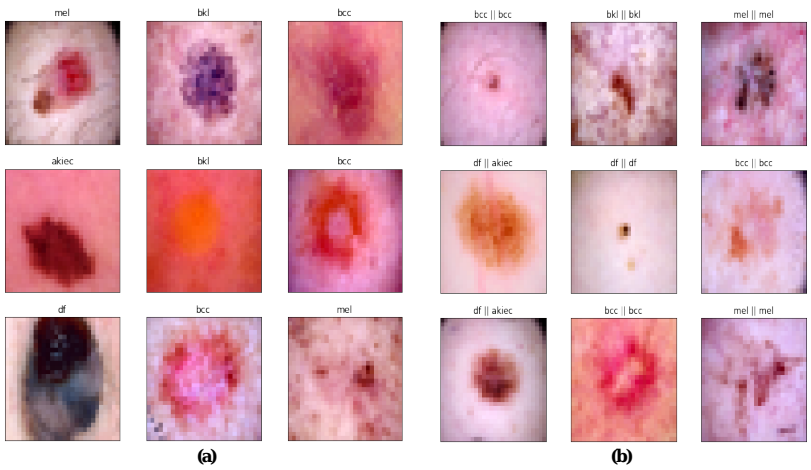


**Fig. 4.** Skin lesion classification

The proposed YOLOv3-DCNN being tested for categorization of benign as well as malignant skin lesions on the MNIST HAM10000 dataset. The characteristics obtained from GLCM, RCT, and CLCM are integrated with YOLO-v3 features which are depicted in figure 4(a). Then, the extracted features are classified using the proposed DCNN approach, which is shown in Figure 4(b). Furthermore, our results indicate that our proposed approach can detect skin lesions earlier and with greater accuracy.

## Performance Parameters

This section describes the performance of our proposed technique whereas; various parameters such as accuracy, precision were employed to assess the effectiveness of the unique technique to identifying as well as categorizing skin lesions.

### *Accuracy*

The most basic intuitive performance metric is accuracy, which seems to be the ratio of precisely predicted observations to all observations. The accuracy is formulated as follows (13)

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \tag{13}$$

Where, TP- True Positive, TN-True Negative, FP- False Positive, FN-False Negative

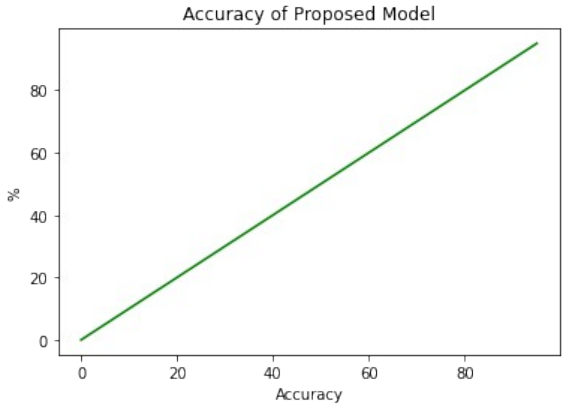


**Fig. 5.** Accuracy of the proposed model

Figure 5 depicts the accuracy of our proposed skin lesion detection approach. When YOLO-V3 characteristics were integrated with GLCM, RCT, as well as CLCM features, our proposed architecture's accuracy improves. The accuracy of our suggested technique is 95%, which describes the effectiveness of the feature fusion approach.

***Precision***

Precision has been described as the ratio of correctly predicted positive observations to total predicted positive observations. The precision is calculated by equation (14),

$$\text{Precision} = \frac{TP}{TP+FP} \tag{14}$$

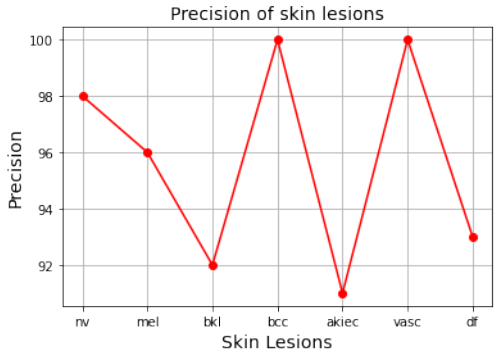


**Fig. 6.** Precision of the proposed model

The precision of the proposed skin lesion categorization approach is shown in figure 6. The DCNN technique improves the precision of the proposed architecture. The precision of our suggested technique attains the highest value of 100% and the lowest value of 91%, which describes the effectiveness of the classification approach.

***Training accuracy and validation accuracy of the proposed model***

Figure 7 depicts the training as well as validation accuracy of our proposed approach. In figure 7, green color graphical lines illustrate the training accuracy, whereas brown color graphical lines illustrate validation accuracy.

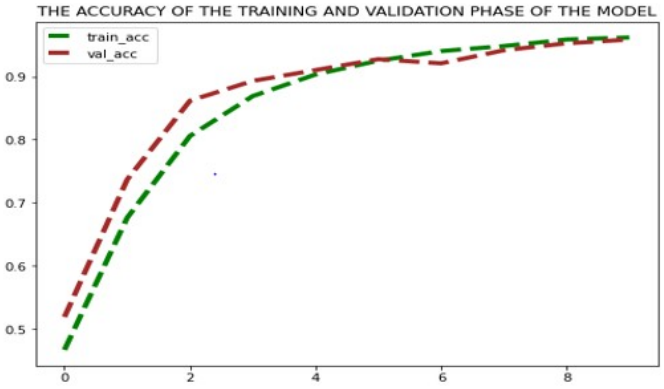


**Fig. 7.** The proposed model's training as well as validation accuracy

The accuracy of training and validation is increasing over time. Our proposed model reaches 0.97 training accuracy and 0.975 validation accuracy over 9 epochs. Furthermore, the training accuracy exceeds the validation accuracy.

***The proposed model's training as well as validation losses***

Figure 8 depicts the training as well as validation losses of our proposed model. In figure 8, orange color graphical lines illustrate the training loss, whereas violet color graphical lines illustrate validation loss.

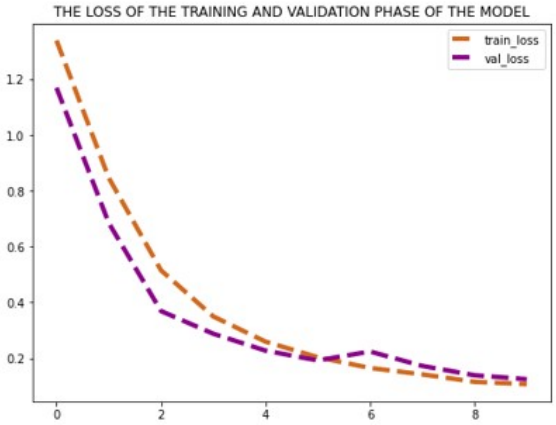


**Fig. 8.** The proposed model's training as well as validation losses

Training as well as validation loss gradually reducing with time. Figure 8 shows that the proposed model obtained less than 0.1 train loss in 9 epochs. In 9.5 epochs, the proposed model obtained less than 0.1 validation loss. Furthermore, the training loss is less severe than the validation loss.

***Confusion Matrix***

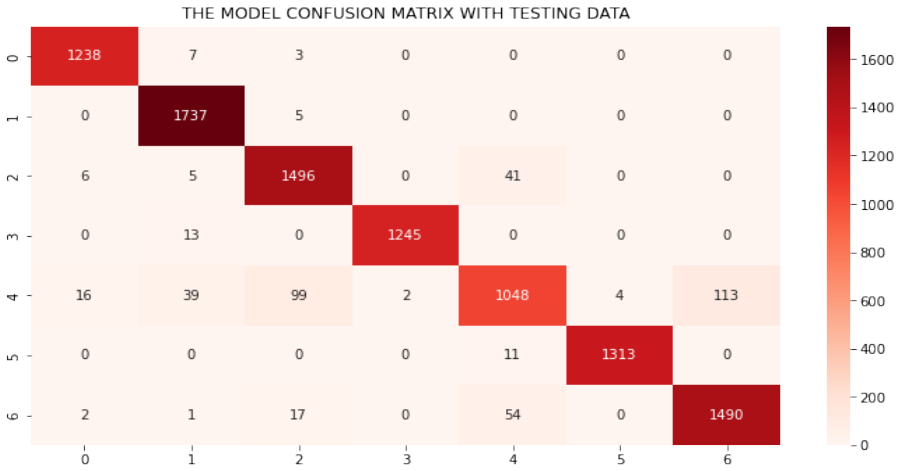


**Fig. 9.** Confusion matrix

Figure 9 displays the skin lesion recognition confusion matrix. The proposed model produces the confusion matrix, in which the diagonal values indicate the count of tuples properly identified by the models and the off-diagonal values denotes the count of tuples erroneously categorized by the models. The better the performance, the higher the diagonal value.

**Comparision Results**

This section explains the proposed technique's comparative outcomes, in which our novel technique is compared to the baseline approach, such as Grey Level Co-occurrence Matrix (GLCM) with Gabor and Color Level Co-occurrence Matrix (CLCM) [37], Grey Level Co-occurrence Matrix (GLCM) with Gabor features [37], Color Level Co-occurrence Matrix (CLCM) with Gabor features [37], and Gabor with Color Level Co-occurrence Matrix (CLCM) [37].

**Table 1.** Overall Accuracy

| **Methods** | **Accuracy (%)** |
|---|---|
| GLCM + Gabor + CLCM [37] | 94 |
| GLCM + Gabor [37] | 90 |
| CLCM + Gabor [37] | 91 |
| Gabor + CLCM [37] | 88 |
| **GLCM+RCT+CL** | **95** |

| CM (Proposed) | |
|---|---|

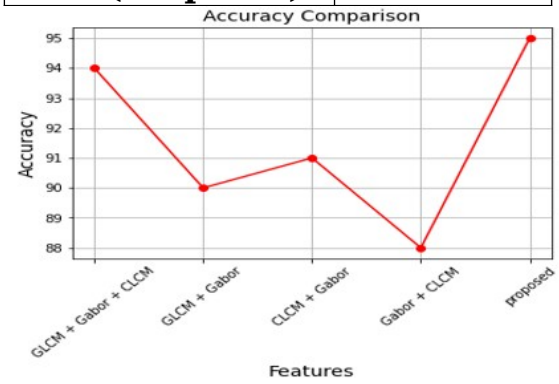


**Fig. 10.** Accuracy Comparison

Figure 10 illustrates the overall comparison of the feature fusion accuracy. The proposed technique attains higher accuracy by incorporating Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform(RCT) and Color moments with QuadHistogram. Our proposed approach compared with the baseline GLCM + Gabor + CLCM [37], GLCM + Gabor [37], CLCM + Gabor, and Gabor + CLCM [37] such as 94%, 90%, 91%, and 88%. As a consequence, our proposed approach has a superior accuracy of 95% than existing methodologies.

**Table 2.** Overall Precision

| Methods | Precision (%) |
|---|---|
| GLCM + Gabor + CLCM [37] | 92 |
| GLCM + Gabor [37] | 87 |
| CLCM + Gabor | 86 |
| Gabor + CLCM | 85 |
| **GLCM+RCT+CLCM (Proposed)** | **94** |

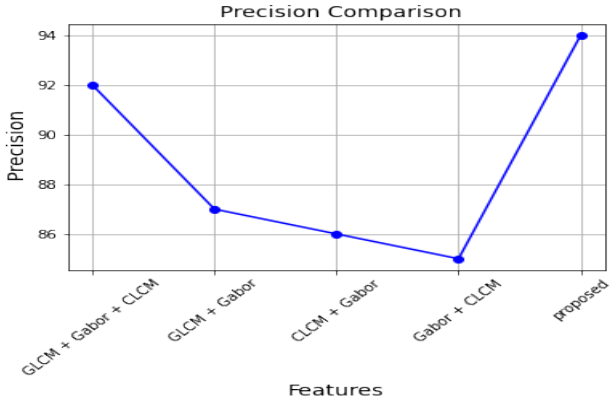

**Fig. 11.** Precision Comparison

Figure 11 depicts the precision for the overall feature fusion comparison. The precision of the proposed technique attains a higher precision by incorporating Grey Level Co-occurrence Matrix (GLCM) with Redundant Contourlet Transform(RCT) and Color moments with QuadHistogram. Our proposed approach compared with the baseline GLCM + Gabor + CLCM [37], GLCM + Gabor [37], CLCM + Gabor, and Gabor + CLCM [37] such as 92%, 87%, 86%, and 85%. As a consequence, our proposed technique has a greater accuracy of 94 % than existing techniques.

**Table 3.** Model Accuracy

| **Methods** | **Accuracy (%)** |
|---|---|
| MobileNet [37] | 80 |
| SSD-MobileNet [37] | 90 |
| Faster R-CNN [37] | 92 |
| **YOLO-v3 - DCNN (Proposed)** | **95** |

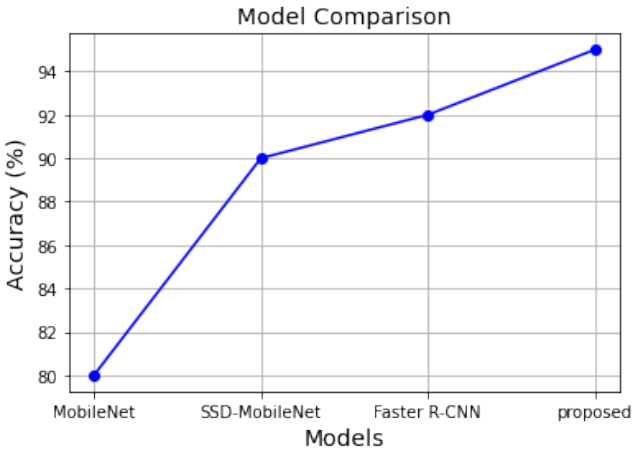


**Fig. 12. Model Comparison**

Figure 12 depicts an overall comparison of model accuracy. The proposed model attains higher accuracy by incorporating YOLO-v3 - DCNN. Our

proposed approach compared with the baseline MobileNet [37], SSD-MobileNet [37], and Faster R-CNN [37] such as 80%, 90%, and 92%. Thus our proposed technique has obtained an accuracy of 95% which is higher than the existing techniques.

## CONCLUSION

Early illness diagnosis is critical in the healthcare and medical fields for treatment planning. Treatment and consequent mitigation were critical for survival if illnesses are recognized early. Also with growth of numerous skin disorders as a result of pollution and other causes, the categorization of surface-level skin cancers as well as diseases is becoming an important subject of study. It is incredibly hard to acquire labelled data that would aid categorization models. Furthermore, low picture quality, fuzzy regions, including noise such as hair or sweat in dermoscopic pictures may have an effect on model training. Moreover, the proposed method obtains color–texture characteristics from input pictures, which are then concatenated with DCNN features for further FCN training, significantly optimizes the classifier's performance. In terms of accuracy and categorization indices, the proposed YOLO v3 - DCNN method outperforms previous models.

## Reference


1. A. L. Byrd, Y. Belkaid, J. A. Segre, The human skin microbiome. *Nat Rev Microbiol* **16**, 143 (2018).
2. D. E. O'Sullivan, D. R. Brenner, P. A. Demers, P. J. Villeneuve, C. M. Friedenreich, W. D. King, N. Zhang, et al., Artificial Intelligence In Medicine 102 (2020) 101756 et al. Indoor tanning and skin cancer in Canada: a meta-analysis and attributable burden estimation. *Cancer Epidemiol* **59**, 1–7 (2019).
3. P. Hylands, Skin cancer: types, diagnosis and prevention. *Heart Fail* **10**, 00 (2019).
4. "Defining cancer". National cancer institute (2014). Archived from the original on 25 June 2014. Retrieved 10 June 2014.
5. M. Elgamal, Automatic Skin Cancer Images Classification. *IJACSA* **4** (2013).
6. Key Statistics for Melanoma Skin Cancer. Am. Cancer Soc. Available online: https://www.cancer.org/content/dam/CRC/PDF/Public/8823.00.pdf (accessed on 8 February 2021).
7. E. Hodis, The somatic genetics of human melanoma. (2018).
8. M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, M. A. Khan, Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer. *IEEE Access* **7**, 90132–90144 (2019).

9. N. Cascinelli , M. Ferrario , T. Tonelli , E. Leo , A possible new tool for clinical diagnosis of melanoma: the computer. *J. Am. Acad. Dermatol*. **16,** 361–367 (1987).

10. B. D. Parameshachari , H. T. Panduranga , S. liberataUllo , Analysis and computation of encryption technique to enhance security of medical images, *IOP Conference Series: Materials Science and Engineering IOP Publishing* **925** (2020).

11. M. Arun , E. Baraneetharan , A. Kanchana , S. Prabu , Detection and monitoring of the asymptotic COVID-19 patients using IoT devices and sensors, *Int. J. Pervas. Comput. Commun*. (2020).

12. FikretErcal, Anurag Chawla, William V. Stoecker, Hsi-Chieh Lee and Randy H. Moss, Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on Biomedical Engineering* **41**(9), 837–845 (1994).

13. M. EmreCelebi, Hassan A. Kingravi, Bakhtiyar Uddin, Hitoshi Iyatomi, Y. Alp Aslandogan, William V. Stoecker and Randy H. Moss, A methodological approach to the classification of dermoscopy images. *Computerized Medical Imaging and Graphics* **31**(6), 362–373 (2007).

14. Paul Wighton, Tim K. Lee, Harvey Lui, David I. McLean and M. Stella Atkins, Generalizing common tasks in automated skin lesion diagnosis. *IEEE Transactions on Information Technology in Biomedicine* **15**(4), 622–629 (2011).

15. IliasMaglogiannis and Konstantinos K. Delibasis, Enhancing classification accuracy utilizing globules and dots features in digital dermoscopy. *Computer Methods and Programs in Biomedicine* **118**(2), 124–133 (2015).

16. CatarinaBarata, M. EmreCelebi and Jorge S. Marques, Improving dermoscopy image classification using color constancy. *IEEE Journal of Biomedical and Health Informatics* **19**(3), 1146–1152 (2014).

17. N. Razmjooy, B. S. Mousavi, F. Soleymani, A real-time mathematical computer method for potato inspection using machine vision. *Comput Math Appl* **63**, 268–79 (2012).

18. V. M. Cohen, E. Pavlidou, J. DaCosta, A. K. Arora, T. Szyszko, M. S. Sagoo, et al., Staginguveal melanoma with whole-body positron-emission tomography/computed tomography and abdominal ultrasound: low incidence of metastatic disease, high incidence of second primary cancers. *Middle East Afr J Ophthalmol* **25**, 91 (2018).

19. KhodaeiHossein, et al., Fuzzy-based heat and power hub models for cost-emission operation of an industrial consumer using compromise programming. *ApplThermEng* **137**, 395–405 (2018).

20. A. Kulkarni, D. Mukhopadhyay, SVM classifier based melanoma image classification. *Res J Pharm Technol* **10**, 4391–2 (2017).

21. K. Narasimhan, V. Elamaran, Wavelet-based energy features for diagnosis of melanoma from dermoscopic images. *Int J Biomed EngTechnol* **20**, 243–52 (2016).

22. H. Rashid, M. A. Tanveer, H. Aqeel Khan, Skin Lesion Classification Using GAN Based Data Augmentation. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany,23–27, 916–919 (July 2019).

23. D. Bisla, A. Choromanska, J. A. Stein, D. Polsky, R. Berman, Towards Automated Melanoma Detection with Deep Learning:Data Purification and Augmentation. arXiv2019, arXiv:1902.06061. Available online: http://arxiv.org/abs/1902.06061 (accessedon 10 February 2021).

24. A. Farag, L. Lu, H. R. Roth, J. Liu, E. Turkbey, R. M. Summers, A Bottom-Up Approach for Pancreas Segmentation UsingCascaded Superpixels and (Deep) Image Patch Labeling. *IEEE Trans. Image Process* **26** 386–399 (2017).

25. F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, A. Bovik, Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model. *IEEE Trans. Med. Imaging* **36**, 849–858 (2017).

26. L. Yu, H. Chen, Q. Dou, J. Qin, P. A. Heng, Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks. *IEEE Trans. Med. Imaging* **36**, 994–1004 (2017).

27. Dorj, Ulzii-Orshikh, Keun-Kwang Lee, Jae-Young Choi and Malrey Lee, The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications* **77**(8), 9909-9924 (2018).

28. Nida, Nudrat, AunIrtaza, Ali Javed, Muhammad HaroonYousaf and Muhammad Tariq Mahmood, Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. *International journal of medical informatics* **124,** 37-48 (2019).

29. Adegun, A. Adekanmi and SerestinaViriri, FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. *IEEE Access* **8** 150377-150396 (2020).

30. Akram, Tallha, Hafiz M. JunaidLodhi, Syed Rameez Naqvi, Sidra Naeem, MajedAlhaisoni, Muhammad Ali, Sajjad Ali Haider and Nadia N. Qadri, A multilevel features selection framework for skin lesion classification. *Human-centric Computing and Information Sciences* **10**(1), 1-26 (2020).

31. Khan, M. Attique, TallhaAkram, Muhammad Sharif, KashifJaved, Muhammad Rashid and Syed Ahmad Chan Bukhari, An integrated framework of skin lesion detection and recognition through saliency method and optimal deep neural network features selection. *Neural Computing and Applications* **32**(20), 15929-15948 (2020).

32. Khan, Muhammad Attique, Muhammad Sharif, TallhaAkram, Syed Ahmad Chan Bukhari and Ramesh Sunder Nayak, Developed Newton-Raphson based deep features selection framework for skin lesion recognition, *Pattern Recognition Letters* **129**, 293-303 (2020).

33. Khan, Muhammad Attique, TallhaAkram, Yu-Dong Zhang and Muhammad Sharif, Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework. *Pattern Recognition Letters* **143,** 58-66 (2021).

34. Sikkandar, Mohamed Yacin, Bader AwadhAlrasheadi, N. B. Prakash, G. R. Hemalakshmi, A. Mohanarathinam and K. Shankar, Deep learning based an automated skin lesion segmentation and intelligent classification model. *Journal of ambient intelligence and humanized computing* **12**(3), 3245-3255 (2021).

35. Saeed, Jwan and SubhiZeebaree, Skin lesion classification based on deep convolutional neural networks architectures, *Journal of Applied Science and Technology Trends* **2**(01), 41-51 (2021).

36. P. Tschandl, C. Rosendahl & H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **5**, 180161 (2018). doi: 10.1038/sdata.2018.161

37. Nersisson, Ruban, Tharun J. Iyer, Alex Noel Joseph Raj and Vijayarajan Rajangam, A Dermoscopic Skin Lesion Classification Technique Using YOLO-CNN and Traditional Feature Model, *Arabian Journal for Science and Engineering* **46**(10), 9797-9808 (2021).